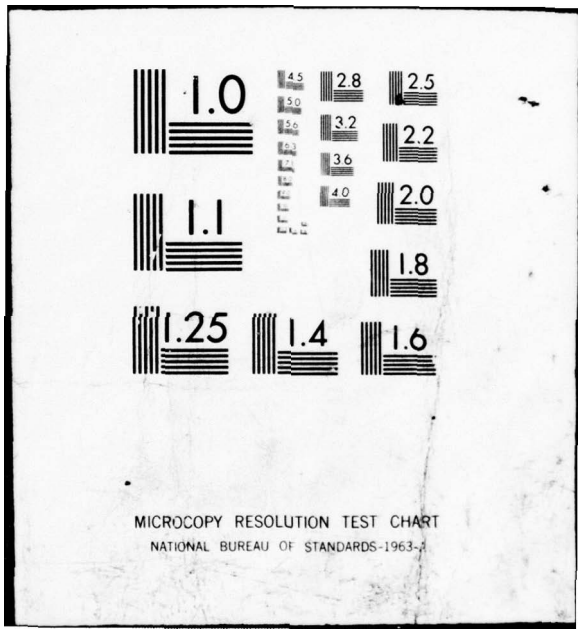
AD-A052 512 MASSACHUSETTS INST OF TECH CAMBRIDGE ARTIFICIAL INTE--ETC F/G 9/4
USING SYNTHETIC IMAGES TO REGISTER REAL IMAGES WITH SURFACE MOD--ETC(U)
AUG 77 B K HORN, B L BACHMAN N00014-75-C-0643
UNCLASSIFIED AI-M-437 NL

| OF |
AD
A062512

END
DATE
FILMED
5-78
DDC


1.0
1.1
1.25
1.4
1.6
1.8
2.0
2.2
2.5
2.8
3.2
3.6
4.0
4.5
5.0
5.6
MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

UNCLASSIFIED
SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER: AIM 437
2. GOVT ACCESSION NO.: AI-M-437
3. RECIPIENT'S CATALOG NUMBER:

4. TITLE (and Subtitle): Using Synthetic Images to Register Real Images with Surface Models
5. TYPE OF REPORT & PERIOD COVERED: MEMO
6. PERFORMING ORG. REPORT NUMBER:

7. AUTHOR(s): Berthold K.P. Horn, Brett L. Bachman
8. CONTRACT OR GRANT NUMBER(s): N00014-75-C-0643

9. PERFORMING ORGANIZATION NAME AND ADDRESS:
Artificial Intelligence Laboratory
545 Technology Square
Cambridge, Massachusetts 02139

10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS:

11. CONTROLLING OFFICE NAME AND ADDRESS:
Advanced Research Projects Agency
1400 Wilson Blvd
Arlington, Virginia 22209

12. REPORT DATE: Aug 77
13. NUMBER OF PAGES: 50

14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office):
Office of Naval Research
Information Systems
Arlington, Virginia 22217

15. SECURITY CLASS. (of this report): UNCLASSIFIED
15a. DECLASSIFICATION/DOWNGRADING SCHEDULE:

52 p.

16. DISTRIBUTION STATEMENT (of this Report)

Distribution of this document is unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

None

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Image Matching
Synthetic Images
Surface models
Albedo maps
Change detection
Terrain classification
Hill-shading
Digital terrain models
Image registration

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

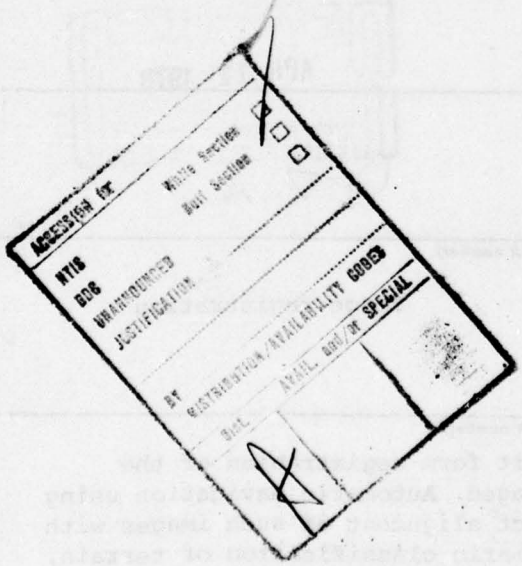
A number of image analysis tasks can benefit form registration of the iamge with a model of the surface being imaged. Automatic navigation using visible light or radar images requires exact alignment of such images with digital terrain models. In addition, automatic classification of terrain, using satellite imagery, requires such alignment to deal correctly with the effects of varying sun angle and surface slope. Even inspection techniques for certain industrial parts may be improved by this means. (cont'd

→ next page

407 483

BLOCK 20 CONCLUDED

~~We achieve~~ is achieved the required alignment by matching the real image with the synthetic image obtained from a surface model and known positions of the light sources. The synthetic image intensity is calculated using the reflectance map, a convenient way of describing surface reflection as a function of surface gradient. ~~We illustrate~~ is illustrated the technique using LANDSAT images and digital terrain models.



MASSACHUSETTS INSTITUTE OF TECHNOLOGY

ARTIFICIAL INTELLIGENCE LABORATORY

A. I. Memo 437 August 1977

USING SYNTHETIC IMAGES TO REGISTER REAL IMAGES WITH SURFACE MODELS

Berthold K. P. Horn and Brett L. Bachman

Abstract. A number of image analysis tasks can benefit from registration of the image with a model of the surface being imaged. Automatic navigation using visible light or radar images requires exact alignment of such images with digital terrain models. In addition, automatic classification of terrain, using satellite imagery, requires such alignment to deal correctly with the effects of varying sun angle and surface slope. Even inspection techniques for certain industrial parts may be improved by this means.

We achieve the required alignment by matching the real image with the synthetic image obtained from a surface model and known positions of the light sources. The synthetic image intensity is calculated using the reflectance map, a convenient way of describing the surface reflection as a function of surface gradient. We illustrate the technique using LANDSAT images and digital terrain models.

This report describes research done at the Artificial Intelligence Laboratory of the Massachusetts Institute of Technology. Support for the laboratory's artificial intelligence research is provided in part by the Advanced Research Projects Agency of the Department of Defense under Office of Naval Research contract N00014-75-C-0643.

CONTENTS.

		page
1.	Motivation	2
2.	Possible Approaches	3
3.	The Reflectance Map	5
4.	Digital Terrain Models	6
5.	The Gradient	8
6.	Position of the Light Source	9
7.	Reflectance as a Function of Gradient	10
8.	Synthetic Images	12
9.	The Real Image	13
10.	Transformation Parameters	14
11.	Choice of Similarity Measure	15
12.	Interpolation Scheme	17
13.	Choice of Normalization Method	18
14.	Locating the Best Match	19
15.	Using Reduced Images	20
16.	Results of Registration Experiments	22
17.	Reasons for Remaining Intensity Mismatches	25
18.	Further Improvement of the Synthetic Image	27
19.	The Influence of Sun Elevation	29
20.	Summary and Conclusions	31

1. Motivation.

Interesting and useful new image analysis methods may be developed if registered image intensity and surface slope information is available. Automatic change detection, for example, seems unattainable without an ability to deal with variations of appearance with changes in the sun's position. In turn, these variations can be understood only in terms of surface topography and reflectance models. Similarly, human cartographers consult both aerial photographs and topographic maps of a region to establish the location of streamlines. Automatic analysis of either of these information sources alone is unlikely to lead to robust methods for performing this task.

Accurate alignment of images with surface models is therefore an important prerequisite for many image understanding tasks. We describe here an automatic method of potentially high accuracy that does not depend on feature extraction or other sophisticated image analysis methods. Instead, all that is required is careful matching of the real with a synthetic image. Because this is an area-based process, it has the potential for sub-pixel accuracy -- accuracy not attainable with techniques dependent on alignment of linear features such as edges or curves. The method is illustrated by registering LANDSAT images with digital terrain models.

2. POSSIBLE APPROACHES.

One way to align a real image with a surface model might be through the use of a reference image obtained under controlled conditions. New images could then be matched against the reference image to achieve alignment. Unfortunately, the appearance of a surface depends quite dramatically on the position of the light source (as seen in figure 1, for example), so that this method works only for a limited daily interval for a limited number of days each year [1]. This problem disappears when one uses synthetic images, since the position of the source can be taken into account.

A more sophisticated process would not match images directly, but first perform a feature extraction process on the real image and then match these features with those found in the reference image. One finds, however, that different features will be seen when lighting changes: for example, ridges and valleys parallel to the illumination direction tend to disappear (see figure 1 again). In addition, the apparent position of a feature as well as its shape may depend somewhat on illumination. More serious may be the present feature extraction schemes' computational cost and lack of robustness. Finally, we should note that the accuracy obtained by matching linear features is likely to be lower than that obtainable with a method based on an aerial match.

One might next consider calculating the shape of the surface from intensities in the image [2]. This, however, is computationally expensive and not likely to be very accurate in view of the variation in the nature of surface cover. A more accurate method, estimating the local gradient using similar methods [3] and then matching these with gradients stored in the terrain model, still involves a great deal of computation.

The method chosen here depends instead on matching the real image with a synthetic image produced from the terrain model. The similarity of the two images depends in part upon how closely the assumed reflectance matches the real one. For mountainous terrain and for images taken with low sun elevations, rather simple assumptions about the reflectance properties of the surface gave very good results. Since all LANDSAT images are taken at about 9:30 local solar time, the sun elevations in this case are fairly small and image registration for all but flat terrain is straightforward.

This implies that LANDSAT images are actually not optimal for automatic terrain classification, since the intensity fluctuations due to varying surface gradients often swamp the intensity fluctuations due to variations in surface cover. An important application of our technique in fact is the removal of the intensity fluctuations due to variations in surface gradient from satellite images in order to facilitate the automatic classification of terrain. To do this, we must model the way the surface reflects light.

3. THE REFLECTANCE MAP.

Work on image understanding has led to a need to model the image-formation process. One aspect of this concerns the geometry of projection, that is, the relationship between the position of a point and the coordinates of its image. Less well understood is the problem of determining image intensities, which requires modelling of the way surfaces reflect light. For a particular kind of surface and a particular placement of light sources, surface reflectance can be plotted as a function of surface gradient (magnitude and direction of slope). The result is called a reflectance map and is usually presented as a contour map of constant reflectance in gradient space [3].

One use of the reflectance map is in the determination of surface shape from intensities [2] in a single image; here, however, it will be employed only in order to generate synthetic images from digital terrain models.

4. DIGITAL TERRAIN MODELS.

Work on computer-based methods for cartography, prediction of side-looking radar imagery for flight-simulators, automatic hill-shading and machines that analyze stereo aerial photography has led to the development of digital terrain models. These models are usually in the form of an array of terrain elevations, z_{ij}, on a square grid.

Data used for this paper's illustrations was entered into a computer after manual interpolation from a contour map and has been used previously in work on automatic hill-shading [4, 5]. It consists of an array of 175 x 240 elevations on a 100-meter grid corresponding to a 17.5 km by 24 km region of Switzerland lying between 7°1' East to 7°15' East and 46°8.5' North to 46°21.5' North. The vertical quantization is 10 meters, and elevations range from 410 meters (in the Rhone valley) to 3210 meters (on the Sommet des Diablerets. The topographic maps used in the generation of the data are "Les Diablerets" (No. 1285) and "Dent de Morcles" (No. 1305), both on a 1:25 000 scale [6]. Extensive data editing was necessary to remove entry errors; some minor distortions of elevations may have resulted.

Manually-entered models of two regions in Canada have also been used [5, 7]. Another data set, covering a region of California, was produced by a digital simulator of a proposed automatic stereo scanner. (Output of two experimental automatic stereo scanners, one built at ETL [8] and one built at RADC [9], could not be obtained.)

The United States Geological Survey [10] supplies digital terrain models on magnetic tape, each covering one square degree of the United States, with a grid spacing of about 208 feet (63.5 m). These models apparently were produced by interpolation from hand-traced contours on existing topographic maps

of the 1:250 000 series. Interpolation to a resolution of .01 inch (0.254 mm) on the original maps fills in elevations between the contours spaced 200 feet (60.96 m) vertically. The final result is smoothed and "generalized" to a considerable extent; nevertheless, this is the most prolific source of surface models available to the public.

5. THE GRADIENT.

A gradient has two components, namely the surface slope along two mutually perpendicular directions. If the surface height, z, is expressed as a function of two coordinates x and y, we define the two components, p and q, of the gradient as the partial derivatives of z with respect to x and y respectively. In particular, a Cartesian coordinate system is erected with the x-axis pointing east, the y-axis north and the z-axis up. Then, p is the slope of the surface in the west-to-east direction, while q is the slope in the south-to-north direction:

$$p = \frac{\partial z}{\partial x} \qquad\qquad q = \frac{\partial z}{\partial y}$$

One can estimate the gradient from the digital terrain model using first differences,

$$p \simeq [z_{(i+1)j} - z_{ij}]/\Delta \qquad q \simeq [z_{i(j+1)} - z_{ij}]/\Delta$$

where Δ is the grid-spacing. More sophisticated schemes are possible [5] for estimating the surface gradient, but are unnecessary.

6. POSITION OF THE LIGHT SOURCES.

In order to be able to calculate the reflectance map, it is necessary to know the location of the light source. In our case the primary source is the sun, and its location can be determined easily by using tables intended for celestial navigation [11, 12, 13] or by straightforward computations [14, 15, 16, 17]. In either case, given the date and time, the azimuth (θ) and the elevation (ϕ) of the sun can be found. Here, azimuth is measured clockwise from North, while elevation is simply the angle between the sun and the horizon (see figure 2). Now one can erect a unit vector at the origin of the coordinate system pointing at the light source,

$$\hat{n}_s = [\sin(\theta)\cos(\phi), \cos(\theta)\cos(\phi), \sin(\phi)].$$

Since a surface element with gradient (p,q) has a normal vector $n = (-p, -q, 1)$, we can identify a particular surface element that happens to be perpendicular to the direction towards the light source. Such a surface element will have a surface normal $n_s = (-p_s, -q_s, 1)$, where $p_s = \sin(\theta)\cot(\phi)$ and $q_s = \cos(\theta)\cot(\phi)$. We can use the gradient (p_s, q_s) as an alternate means of specifying the position of the source.

In work on automatic hill-shading, for example, one uses $p_s = -0.707$ and $q_s = 0.707$ to agree with standard cartographic conventions which require that the light source be in the North-west at 45° elevation ($\theta = (7/4)\pi$, $\phi = \pi/4$) [5].

7. REFLECTANCE AS A FUNCTION OF THE GRADIENT.

Reflectance of a surface can be expressed as a function of the incident angle (i), the emittance angle (e) and the phase angle (g) (see figure 3). We use a simple, idealized reflectance model for the surface material,

$$\Phi_1(i, e, g) = \rho \cos(i).$$

This reflectance function models a surface which, as a perfect diffuser, appears equally bright from all viewing directions. Here, ρ is an "albedo" factor and the cosine of the incident angle simply accounts for the foreshortening of the surface element as seen from the source. More sophisticated models of surface reflectance are possible [3], but are unnecessary for this application.

The incident angle is the angle between the local normal $(-p, -q, 1)$ and the direction to the light source $(-p_s, -q_s, 1)$. The cosine of this angle can then be found by taking the dot-product of the corresponding unit vectors,

$$\cos(i) = \frac{(1 + p_s p + q_s q)}{\sqrt{1 + p_s^2 + q_s^2}\sqrt{1 + p^2 + q^2}}$$

Finally,

$$\Phi_1(p,q) = \frac{\rho\,(1 + p_s p + q_s q)}{\sqrt{1 + p_s^2 + q_s^2}\sqrt{1 + p^2 + q^2}}$$

Another reflectance function, similar to that of materials in the maria of the moon and rocky planets [2, 18], is a little easier to calculate:

$$\Phi_2(p,q) = \rho \cos(i)/\cos(e) = \frac{\rho\,(1 + p_s p + q_s q)}{\sqrt{1 + p_s^2 + q_s^2}}$$

This reflectance function models a surface which reflects equal amounts of light in all directions. For small slopes and low sun elevations, it is very much like the first one, since then $(1 + p^2 + q^2)$ will be near unity. Both functions were tried and both produce good alignment -- in fact, it is difficult to distinguish synthetic images produced using these two reflectance functions.

8. SYNTHETIC IMAGES.

Given the projection equations that relate points on the objects to images of said points, and given a terrain model allowing calculation of surface gradient, it is possible to predict how an image would look under given illuminating conditions, provided the reflectance map is available. We assume simple orthographic projection here as appropriate for a distant spacecraft looking vertically down with a narrow angle of view. Perspective projection would require a few minor changes in the algorithm.

The process of producing the synthetic image is simple. An estimate of the gradient is made for each point in the digital terrain model by considering neighboring elevations. The gradient's components, p and q, are then used to look up or calculate the expected reflectance. An appropriate intensity is placed in the image at the point determined by the projection equation. All computations are simple and local, and the work grows linearly with the number of picture cells in the synthetic image.

Sample synthetic images are shown in figure 1. The two images are of the same region with differences in assumed location of the light source. In figure 2a the sun is at an elevation of 34° and azimuth of 153°, corresponding to its true position at 9:52 G.M.T., 1972/Oct/9, while for figure 2b it was at an elevation of 28° and an azimuth of 223°, corresponding to its position at 13:48 G.M.T. later on the same day. The corresponding reflectance maps are shown in figure 4.

Reflectance maps for the simpler reflectance function $\Phi_2(p,q)$ under the same circumstances are shown in figure 5. Note that near the origin there is very little difference between $\Phi_1(p,q)$ and $\Phi_2(p,q)$. Since most surface elements in this terrain model have slopes less than $1/\sqrt{2}$, as shown in the

scattergram (see figure 6), synthetic images produced using these two reflectance maps are similar.

Since the elevation data is typically rather coarsely quantized as a result of the fixed contour intervals on the base map, p and q usually take on only a few discrete values. In this case, it is convenient to establish a lookup table for the reflectance map by simply precalculating the reflectance for these values. Models with arbitrarily complex reflectance functions can then be easily accomodated as can reflectance functions determined experimentally and known only for a discrete set of surface orientations.

Since the real image was somewhat smoothed in the process of being reproduced and digitized, we found it advantageous to perform a similar smoothing operation of the synthetic images so that the resolution of the two approximately matched. Alignment of real and synthetic images was, however, not dependent on this refinement.

9. THE REAL IMAGE.

The image used for this paper's illustrations is a portion of a LANDSAT [1, 19] image acquired about 9:52 G.M.T. 1972/Oct/9 (ERTS-1 1078-09555). We used channel 6 (near infra-red, .7μ to .8μ), although all four channels (4, 5, 6, & 7) appeared suitable -- with channel 4 (green, .5μ to .6μ) being most sensitive to water in the air column between the satellite and the ground, and channel 7 best at penetrating even thin layers of clouds and snow. Figure 7 compares an enlargement of the original transparency with the synthetic image generated from the terrain model.

A slow-scan vidicon camera (Spatial Data Systems 108) was used to digitize the positive transparency of 1:1 000 000 scale. Individual picture cells were about .1 mm on a side in order to match roughly the resolution of the synthetic image data. In recent work, we used a more accurate version digitized on a drum-scanner (Optronics Photoscan 1000), again with a .1 mm resolution on the film. Note that the "footprint" of a LANDSAT picture cell is about 79 x 79 meters [1], compatible with the resolution of typical digital terrain models. The digitized image used for the illustrations in this paper is of lower resolution, however, due to limitations of the optics and electron-optics of the digitizing system. In future studies we intend to use the computer-compatible tapes supplied by EROS [19].

Alignment of real images with terrain models is possible even with low quality image data, but terrain classification using the aligned image and digital surface model requires high quality data.

We generated image output, as for figures 1a, 1b, 7a, and 11, on a drum film-writer (Optronics Photowriter 1500) and interpolated to alleviate undesirable raster effects due to the relatively small number of picture cells in each image.

10. TRANSFORMATION PARAMETERS.

Before we can match the synthetic and the real image, we must determine the nature of the transformation between them. If the real image truly is an orthographic projection obtained by looking straight down, it is possible to describe this transformation as a combination of a translation, a rotation and a scale change. If we use x and y to designate points in the synthetic image and x' and y' for points in the real image, we may write:

$$\begin{bmatrix} x' - x'_o \\ y' - y'_o \end{bmatrix} = s \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} x - x_o \\ y - y_o \end{bmatrix} + \begin{bmatrix} \Delta x \\ \Delta y \end{bmatrix}$$

where Δx and Δy are the shifts in x' and y' respectively, θ is the angle of rotation and s is the scale factor. Rotation and scaling take place relative to the centers (x_o, y_o) and (x'_o, y'_o) of the two images in order to better decouple the effects of rotation and scaling from translations. That is, the average shift in x' and y' induced by a change in rotation angle or scale is zero.

In our case, the available terrain model restricts in size the synthetic image. The area over which matching of the two will be performed is thus always fixed by the border of the synthetic image. The geometry of the coordinate transformation is illustrated in figure 8.

11. CHOICE OF SIMILARITY MEASURE.

In order to determine the best set of transformation parameters (Δx, Δy, s, θ), one must be able to measure how closely the images match for a particular choice of parameter values. Let S_{ij} be the intensity of the synthetic image at the i^{th} picture cell across in the j^{th} row from the bottom of the image, and define R_{ij} similarly for the real image. Because of the nature of the coordinate transformation, we cannot expect that the point in the real image corresponding to the point (i,j) in the synthetic image will fall precisely on one of the picture cells. Consequently, S_{ij} will have to be compared with R(x',y'), which is interpolated from the array of real image intensities. Here (x',y') is obtained from (i,j) by the transformation described in the previous section.

One measure of difference between the two images may be obtained by summing the absolute values of differences over the whole array. Alternately, one might sum the squares of the differences:

$$\sum_{i=1}^{n} \sum_{j=1}^{m} \{S_{ij} - R(x',y')\}^2$$

This measure will be <u>minimal</u> for exact alignment of the images. Expanding the square, one decomposes this result into three terms, the first being the sum of S_{ij}^2, the last the sum of $R^2(x',y')$. The first is constant, since we always use the full synthetic image; the last varies slowly as different regions of the real image are covered. The sum of $S_{ij}R(x',y')$ is interesting since this term varies most rapidly with changes in the transformation. In fact, a very useful measure of the similarity of the two images is the correlation:

$$\sum_{i=1}^{n} \sum_{j=1}^{m} S_{ij} R(x',y').$$

This measure will be <u>maximal</u> when the images are properly aligned. It has the advantage of being relatively insensitive to constant multiplying factors. These may arise in the real image due to changes in the adjustment of the optical or electronic systems.

Note that image intensity is the product of a constant factor which depends on the details of the imaging system (such as the lens opening and the focal length), the intensity of the illumination striking the surface, and the reflectance of the surface. We assume all but the last factor is constant and thus speak interchangeably of changes in surface reflectance and changes in image intensities.

12. INTERPOLATION SCHEME.

The real image intensity at the point (x',y') has to be estimated from the array of known image intensities. If we let $k = \lfloor x' \rfloor$, and $\ell = \lfloor y' \rfloor$ be the integer parts of x' and y', then R(x',y') can be estimated from $R_{k\ell}$, $R_{(k+1)\ell}$, $R_{k(\ell+1)}$ and $R_{(k+1)(\ell+1)}$ by linear interpolation (see figure 9).

$$R_{\ell}(x') = (k + 1 - x')R_{k\ell} + (x' - k)R_{(k+1)\ell}$$

$$R_{(\ell+1)}(x') = (k + 1 - x')R_{k(\ell+1)} + (x' - k)R_{(k+1)(\ell+1)}$$

$$R(x',y') = (\ell + 1 - y')R_{\ell}(x') + (y' - \ell)R_{(\ell+1)}(x')$$

The answer is independent of the order of interpolation and, in fact, corresponds to the result obtained by fitting a polynomial of the form $(a + bx' + xy' + dx'y')$ to the values at the four indicated points. Alignment is not impaired, however, when nearest neighbor interpolation is used instead. This may be a result of the smoothing of the real image as previously described.

13. CHOICE OF NORMALIZATION METHOD.

High output may result as the transformation is changed simply because the region of the real image used happens to have a high average gray-level. Spurious background slopes and false maxima may then result if the raw correlation is used. For this and other reasons, it is convenient to normalize. One approach essentially amounts to dividing each of the two images by its standard deviation; alternately, one can divide the raw correlation by

$$\sqrt{\sum_{i=j}^{n}\sum_{j=1}^{m} S_{ij}^2} \quad x \quad \sqrt{\sum_{i=j}^{n}\sum_{j=1}^{m} R^2(x',y')}$$

An additional advantage is that a perfect match of the two images now corresponds to a normalized correlation of one. An alternate method uses a normalization factor that is slightly easier to compute and which has certain advantages if the standard deviations of the two images are similar. Instead of using the geometric mean, Hans Moravec proposes the arithmetic mean [20].

$$\left[\sum_{i=1}^{n}\sum_{j=1}^{m} S_{ij}^2 + \sum_{i=1}^{n}\sum_{j=1}^{m} R^2(x',y') \right] /2$$

The first term need not be recomputed, since the full synthetic image is always used. Since we found the alignment procedure insensitive to the choice of normalization method, we used the second in our illustrations.

14. LOCATING THE BEST MATCH.

Now that we have shown how to calculate a good similarity measure, we must find a method to find efficiently the best possible transformation parameters. Exhaustive search is clearly out of the question. Fortunately, the similarity measure allows the use of standard hill-climbing techniques. This is because it tends to vary smoothly with changes in parameters and often is monotonic (at least for small ranges of the parameters).

When images are not seriously misaligned, profiles of the similarity measure usually are unimodal with a well-defined peak when plotted against one of the four parameters of the transformation (see figure 10). It is possible to optimize each parameter in turn, using simple search techniques in one dimension. The process can then be iterated. A few passes of this process typically produce convergence. (More sophisticated schemes could reduce the amount of computation, but were not explored).

When the images are initially <u>not</u> reasonably aligned, more care has to be taken to avoid being trapped by local maxima. Solving this problem using more extensive search leads to prohibitively lengthy computations. We need a way of reducing the cost of comparing images.

15. USING REDUCE IMAGES.

One way to reduce the computation is to use only sub-images or "windows" extracted from the original images. This is useful for fine matching, but is not satisfactory here because of the lack of global context.

Alternately, one might use sampled images obtained by picking one image intensity to represent a small block of image intensities. This is satisfactory as long as the original images are smoothed and do not have any high resolution features. If this is not the case, aliasing due to under-sampling will produce images of poor quality unsuitable for comparisons.

One solution to this dilemma is to low-pass filter the images before sampling. A simple approximation to this process uses averages of small blocks of image intensities. The easiest method involves making one image intensity in the reduced image equal to the average of a 2 x 2 block of intensities in the original image. This technique can be applied repeatedly to produce ever smaller images and has been used in a number of other applications [20, 21].

The results of the application of this reduction process to real and synthetic images can be seen in figure 11. First, the most highly reduced image is used to get coarse alignment. In this case extensive search in the parameter space is permissible, since the number of picture cells in the images to be matched is very small. This coarse alignment is then refined using the next larger reduced images (with four times as many picture cells). Finally, the full resolution images are used directly to fine tune the alignment. False local maxima are, fortunately, much rarer with the highly reduced pictures, thus further speeding the search process. It is as if the high resolution features are the ones leading to false local maxima.

We found it best, by the way, to determine good values for the translations first, then rotation and, finally, scale change. Naturally when searching for a peak value as a function of one parameter, the best values found so far for the other parameters are used.

16. RESULTS OF REGISTRATION EXPERIMENTS.

We matched the real and synthetic images using the similarity measure and search technique just described. We tried several combinations of implementation details, and in all cases achieved alignment which corresponded to a high value of the normalized correlation, very close to that determined manually. For the images shown here, the normalized correlation coefficient reaches .92 for optimum alignment, and the match is such that no features are more than two picture cells from the expected place, with almost all closer than one. (The major errors in position appear to be due to perspective distortion, as described later, with which the process is not designed to cope). The accuracy with which translation, rotation and scaling were determined can be estimated from the above statement.

Overall, the process appears quite successful, even with degraded data and over a wide range of choices of implementation details. Details of interpolation, normalization, search technique, and even the reflectance map do not matter a great deal.

Having stated that alignment can be accurately achieved, we may now ask how similar the real and synthetic images are. There are a number of uninformative numerical ways of answering this question. Graphic illustrations, such as images of the differences between the real and synthetic image, are more easily understood. For example, we plot real image intensity versus synthetic image intensity in figure 12. Although one might expect a straight line of slope one, the scattergram shows clusters of points, some near the expected line, some not.

The cluster of points indicated by the arrow labelled A (figure 12) corresponds chiefly to image points showing cloud or snow cover, with intensity

sufficient to saturate the image digitizer. Here the real image intensity exceeds the synthetic image intensity. Arrow B indicates the cluster of points which corresponds to shadowed points. Those near the vertical axis and to its left come from self-shadowed surface elements, while those to the right are regions lying inside shadows cast by other portions of the surface. These cast shadows are not simulated in the synthetic image at the moment. Here the synthetic image is brighter than the real image. Finally, the cluster of points indicated by arrow C arises from the valley floor, which covers a fairly large area and has essentially zero gradient. As a result, the synthetic image has constant intensity here, while the real image shows both darker features (such as the river) and brighter ones (such as those due to the cities and vegetation cover). Most of the ground cover in the valley appears to have higher "albedo" than the bare rock which is exposed in the higher regions, as suggested by the position of this cluster <u>above</u> the line of slope one.

If one were to remove these three clusters of points, the remainder would form one elongated cluster with major axis at about 45°. This shows that, while there may not be an accurate point-by-point equality of intensities, there is a high correlation between intensities in the real and synthetic images.

Note, by the way, that no quantization of intensity is apparent in these scattergrams. This is a result of the smoothing applied to the synthetic image and the interpolation used on the real image. Without smoothing, the synthetic image has fairly coarse quantization levels because of the coarse quantization of elevations as indicated earlier. Without interpolation, the real image, too, has fairly coarse quantization due to the image digitization procedure.

Finally, note that we achieve our goal of obtaining accurate alignment. Detailed matching of synthetic and real image intensity is a new problem which can be approached now that the problem of image registration has been solved.

17. REASONS FOR REMAINING INTENSITY MISMATCHES.

We may need more accurate prediction of image intensities for some applications of aligned image intensity and surface gradient information. Thus, it is useful to analyze the reasons for the differences noted between the synthetic and the real image:

Satellite Imaging. Geometric distortion in satellite imagery may be small but noticeable and traceable to several sources [1]. Shifts of several hundred meters can arise. Perspective distortion for the image used here amounted to about 200 meters on the highest peaks, for example.

Intensity distortions are caused by the fact that scan lines are not all sensed by the same sensor [1]. Electronic noise and atmospheric attenuation, dispersion and scattering are also important for some of the spectral bands.

Digitization. When film transparencies are digitized, the resolution limitations of the optics and the nonlinear response of the film are important. More large errors are introduced if an electron-optic device is used. These typically introduce geometric distortions, nonlinearity and nonuniformity of response. Picture cells may not be square and axes not perpendicular.

Terrain Model. Inaccuracies due to manual entry and editing are common in present day digital terrain models. In addition, the contour maps used commonly as source information are already liberally "generalized" and smoothed by the cartographer. Finally, the estimation of surface gradient is likely to be crude, since the

data in such maps is intended to be accurate in elevation, not in the partial derivatives of elevation. Coarse quantization of the gradient is one effect of this that has already been mentioned. We hope that terrain models produced by automatic stereo comparators in the future will not suffer from all of these shortcomings.

Reflectance. The assumption of uniform reflectance and the modelling of reflectance by means of the simple, rather ad hoc functions used here contribute errors to the synthetic image. More seriously, cast shadows are not modelled. Illumination from the sky and mutual illumination between mountain slopes are less important. Including even crude surface cover information improves the match between the synthetic and the real image.

Water. In its various forms, water can produce large mismatches since, at least for the shorter wavelengths, moisture in the atmosphere contributes to attenuation and scattering of light. In liquid form, water produces bright, obscuring areas in the form of clouds and dark regions such as rivers and lakes. Snow and ice provide highly reflective areas which produce large mismatches.

In view of all these factors, it is surprising that a match as good as that in figure 12 is possible.

18. FURTHER IMPROVEMENT OF THE SYNTHETIC IMAGE.

Using the original digital tapes [19] would eliminate the errors we believe are due to the digitization process. Most of the geometric distortion can be dealt with as well [1]. Further match improvement must come from better synthetic images.

The most significant step here would be the inclusion of surface cover information. Even a coarse categorization into materials of grossly differing "albedo" might be useful. Conversely, of course, one can exploit the difference in intensities between the real and the synthetic image to estimate surface reflectance. Since alignment is possible without accurate reflectance models, the ratio of real to synthetic intensity (a measure akin to albedo) can be used in terrain classification, particularly if it is calculated for each of the spectral bands.

Cast shadows are fairly easy to deal with, if we implement a hidden-surface algorithm to determine which surface elements can be seen from the source. This computation can be done fairly quickly using a well known algorithm [22]. Sky illumination in shadowed areas presents no great stumbling block in this regard.

The quality of terrain models is likely to increase most rapidly when fully automatic scanning stereo comparators become available. Until then, hand-editing of hand-traced information will have to be used to limit the errors in the estimation of gradient.

One notion that shows great promise is that of masks derived from both the terrain model and the real image. The masks are used to limit the correlation operation to those areas which are not as likely to lead to mismatches. Areas of very high intensity in the image, for example, may suggest cloud or

snow cover, and ought not to be used in the matching operation. Similarly, it may be that areas of certain elevations and surface gradients are better than others for matching. The correlation can be improved considerably if we use only those regions above the elevations at which dense vegetation exists and below the elevation at which snow may have accumulated. A slightly more sophisticated method would note that snow tends to remain longer on north-facing slopes.

19. THE INFLUENCE OF SUN ELEVATION.

Aerial or satellite photographs obtained when the sun is low in the sky show the surface topography most clearly. In this case, the surface gradient is the major factor in determining surface reflectance. Ridges and valleys stand out in stark relief, and one gets an immediate impression of the shape of portions of the surface. Conversely, variations in surface cover tend to be most important when the sun is high in the sky. Photographs obtained under such conditions are difficult to align with a topographic map -- at least for a beginner.

What is the sun elevation for which these two effects are about equally important? Finding this value will allow us to separate the imaging situations into two classes: those which are more suited for determining topography and those which are more conducive to terrain classification success. We will use a simple model of surface reflectance. Suppose that the surface has materials varying in "albedo" between ρ_1 and ρ_2. Next, suppose that the surface slopes are all less than or equal to tan(e). The incident angles will vary between $e - (90° - \phi)$ and $e + (90° - \phi)$, where ϕ is the elevation of the sun. If we use the same simple reflectance function employed before, we find that for the two influences on reflectance to be just equal:

$$\rho_1 \cos(e + 90° - \phi) = \rho_2 \cos(e - 90° + \phi)$$

Expanding the cosine and rearranging this equation leads to:

$$\tan(\phi) = \left| \frac{\rho_1 + \rho_2}{\rho_1 - \rho_2} \right| \tan(e)$$

When, for example, the surface materials have reflectances covering a range of two to one and the sun elevation is 35°, then regions with surface slopes above approximately 0.23 (e ≃ 13°) will have image intensities affected more by surface gradient than by surface cover. Conversely, flatter surfaces will result in images more affected by variations in surface cover than by the area's topography.

One possible conclusion is that alignment of images with terrain models is feasible without detailed knowledge of the surface materials if the sun elevation is small and the surface slopes are high. Since LANDSAT images are taken at about 9:30 local solar time [1], the first condition is satisfied and alignment of these images is possible even in only lightly undulating terrain.

. Conversely, if one is attempting terrain classification in anything but flat regions, high sun elevations are needed. Curiously, LANDSAT does not provide such imagery despite the fact that one of its main applications is in land use classification.

20. SUMMARY AND CONCLUSIONS.

We have seen that real images can be aligned with surface models using synthetic images as an intermediary. This process works well despite many factors which contribute to differences between the real and synthetic images. The computations, while lengthy, are straightforward, and reduced images have been used to speed up the search for the best set of transformation parameters.

Several applications of aligned images and surface information have been presented. More can be found; for problems in a different domain, see reference [23] for example. Aside from change detection, passive navigation, photo-interpretation, and inspection of industrial parts, perhaps the most important application lies in the area of terrain classification.

So far, no account has been taken of the effect of varying surface gradient, sun position, and reflective properties of ground cover. Recently, some interest has arisen in an understanding of how surface layers reflect light [24, 25, 26] and how this understanding might aid the interpretation of satellite imagery [27, 28, 29].

It is imperative that interpretation of image information be guided by an understanding of the imaging process. This, in turn, can be achieved if one understands how light is reflected from various surfaces and how this might be affected by such factors as light source position, moisture content and point in the growth cycle of vegetation.

ACKNOWLEDGEMENTS.

We would like to thank Professor Kurt Brassel (State University of New York, Buffalo, New York), Professor Thomas Peucker (Simon Fraser University, Burnaby, B. C., Canada), Dr. Bala (EIKONICS, Burlington, Massachusetts), George E. Lukes (U.S.A. E.T.L., Fort Belvoir, Virginia), and Dr. Harry Barrow (Stanford Research Institute, Palo Alto, California) for generously supplying digital terrain models.

We would also like to express our appreciation to Blenda Horn, Eva Kampits, Karen Prendergast, and Patrick Winston who provided numerous helpful suggestions and improved the presentation of the results. Special thanks are due Al Paone and Russ Clark of the Massachusetts Institute of Technology Graphic Arts Service for help in the preparation of the photographic reproductions.

This report describes research done at the Artificial Intelligence Laboratory of the Massachusetts Institute of Technology. Support for the laboratory's artificial intelligence research is provided in part by the Advanced Research Projects Agency of the Department of Defense under Office of Naval Research contract N00014-75-C-0643.

REFERENCES.

1. R. Bernstein, "Digital image processing of earth observation sensor data," IBM Journal of Research and Development, Vol. 20, pp. 40-57. January 1976.

2. B. K. P. Horn, "Obtaining shape from shading information," in P. H. Winston (ed.), The Psychology of Machine Vision, New York, McGraw-Hill, 1975, pp. 115-155.

3. B. K. P. Horn, "Understanding image intensity," Artificial Intelligence, Vol. 8, pp. 201-231, April 1977.

4. K. Brassel, Modelle und Versuch zur automatischen Schraglicht-schattierung, 7250 Klosters, Switzerland, Buchdruckerei E. Brassel, 1973.

5. B. K. P. Horn, "Automatic hill-shading using the reflectance map," unpublished, 1976.

6. Carte nationale de la Suisse, Swiss National Tourist Office, New York, New York 10020.

7. T. K. Peucker, "Computer cartography," Resource paper no. 17, Association of American Geographers, Washington, D.C. 1972.

8. Bala, "Experimental heterodyne optical correlator," ETL-0071, U.S. Army Engineer Topographic Laboratories, Fort Belvoir, Virginia 22060 (October 1976).

9. "AS-11B-X automated stereo mapper," RADC-TR-76-100, Rome Air Development, Griffiths Air Force Base, New York 13441, April 1976.

10. "N.C.I.C. digital terrain file," and "N.C.I.C. 1:250 000-scale digital terrain library," National Cartographic Information Center, U.S. Geological Survey, Reston, Virginia 22092.

11. "The nautical almanac for the year 1972," U.S. Government Printing Office, Washington, D.C. 20402.

12. "The American ephemeris and nautical almanac for the year 1972," U.S. Government Printing Office, Washington, D.C. 20402.

13. "Explanatory supplement to the astronomical ephemeris and the American ephemeris and nautical almanac," Her Majesty's Stationery Office, London, England, 1961.

14. Simon Newcomb, A Compendium of Spherical Astronomy, New York, McMillan & Co., 1895, 1906.

15. W. M. Smart, Textbook on spherical astronomy, Cambridge, Cambridge University Press, 1931, 1965.

16. E. W. Woolard & G. M. Clemance, Spherical Astronomy, Cambridge, Cambridge University Press, 1966.

17. B. K. P. Horn, "Celestial navigation suite for a programmable calculator," unpublished, 1976.

18. B. Hapke, "An improved theoretical lunar photometric function," The Astronomical Journal, Vol. 71, pp. 333-339.

19. "Computer compatible tape (CCT) index," EROS Data Center, Sioux Falls, South Dakota 57198.

20. H. P. Moravec, "Techniques towards automatic visual obstacle avoidance," Proceedings of the Fifth International Conference on Artificial Intelligence, Cambridge MA, pp. (not yet published).

21. S. L. Tanimoto, "Pictorial feature distortion in a pyramid," Computer Graphics and Image Processing, Vol. 5, pp. 333-352, September 1976.

22. I. E. Sutherland, R. F. Sproull & R. A. Schumacker, "A characterization of ten hidden-surface algorithms," Computing Surveys, Vol. 6, pp. 1-55.

23. D. Nitzan, A. E. Brain & R. O. Duda, "The measurement and use of registered reflectance and range data in scene analysis," Proc. of the IEEE, Vol. 65, pp. 206-220.

24. E. L. Simmons, "Particle model theory of diffuse reflectance: effect of non-uniform particle size," Applied Optics, Vol. 15, pp. 603-604, 1976.

25. W. W. M. Wendlandt & H. G. Hecht, Reflectance Spectroscopy, New York, Interscience Publishers, 1966.

26. P. Oetking, "Photometric studies of diffusely reflecting surfaces with applications to the brightness of the moon," Journal of Geophysical Research, Vol. 71, pp. 2505-2514, May 1966.

27. C. J. Tucker & M. W. Garret, "Leaf optical system modeled as a stochastic process, Applied Optics, Vol. 16, pp. 635-642, March 1977.

28. C. J. Tucker, "Asymptotic nature of grass canopy spectral reflectance," Applied Optics, Vol. 16, pp. 1151-1156, May 1977.

29. M. J. Duggin, "Likely effects of solar elevation on the quantification of changes in vegetation with maturity using sequential LANDSAT images, Applied Optics, Vol. 16, pp. 521-523, March 1977.

FIGURE CAPTIONS.

Figure 1a. Early morning (9:52 G.M.T.) synthetic image.

Figure 1b. Early afternoon (13:48 G.M.T.) synthetic image.

Figure 2. Definition of azimuth and elevation of the sun.

Figure 3. The geometry of light reflection from a surface element is governed by the incident angle, i, the emittance angle, e, and the phase angle, g.

Figure 4a. Reflectance map used in the synthesis of figure 1a. The curves shown are contours of constant $\Phi_1(p,q)$ for $\rho = 1$.

Figure 4b. Reflectance map used in the synthesis of figure 1b.

Figure 5a. Alternate reflectance map, which could have been used in place of the one shown in figure 4a. The curves shown are contours of constant $\Phi_2(p,q)$ for $\rho = 1$.

Figure 6. Scattergram of surface gradients found in the digital terrain model.

Figure 7a. Synthetic image used in the alignment experiments.

Figure 7b. Enlargement of the transparency containing the real image used in the alignment experiments.

Figure 8. Coordinate transformation from synthetic image to real image.

Figure 9. Simple interpolation scheme applied to the real image array.

Figure 10a. Variation of similarity measure with translation in x direction.

Figure 10b. Variation of similarity measure with translation in y direction.

Figure 10c. Variation of similarity measure with rotation.

Figure 10d. Variation of similarity measure with scale changes.

Figure 11. Successive reduction by factors of two applied to both the synthetic (left) and the real (right) image.

Figure 12a. Scattergram of real image intensities versus synthetic image intensities based on $\Phi_1(p,q)$.

Figure 12b. Scattergram of real image intensities versus synthetic image intensities based on $\Phi_2(p,q)$.

LAMBERT DENT_DE_MORCLES

IMAGE2.2 at photowriter resolution 3.
Data linearly interpolated. JULY 14, 1977.

FIGURE 1a

LAMBERT DENT_DE_MORCLES

IMAGE2.4, second exposure. Resolution 3.
data interpolated. Date 15 JULY 1977.

FIGURE 1b

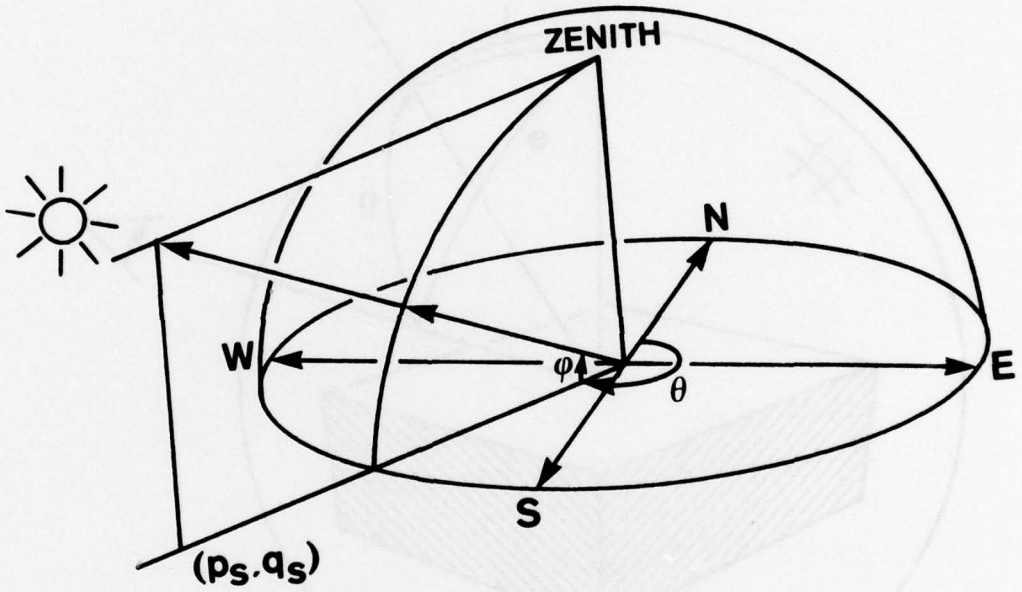


FIGURE 2

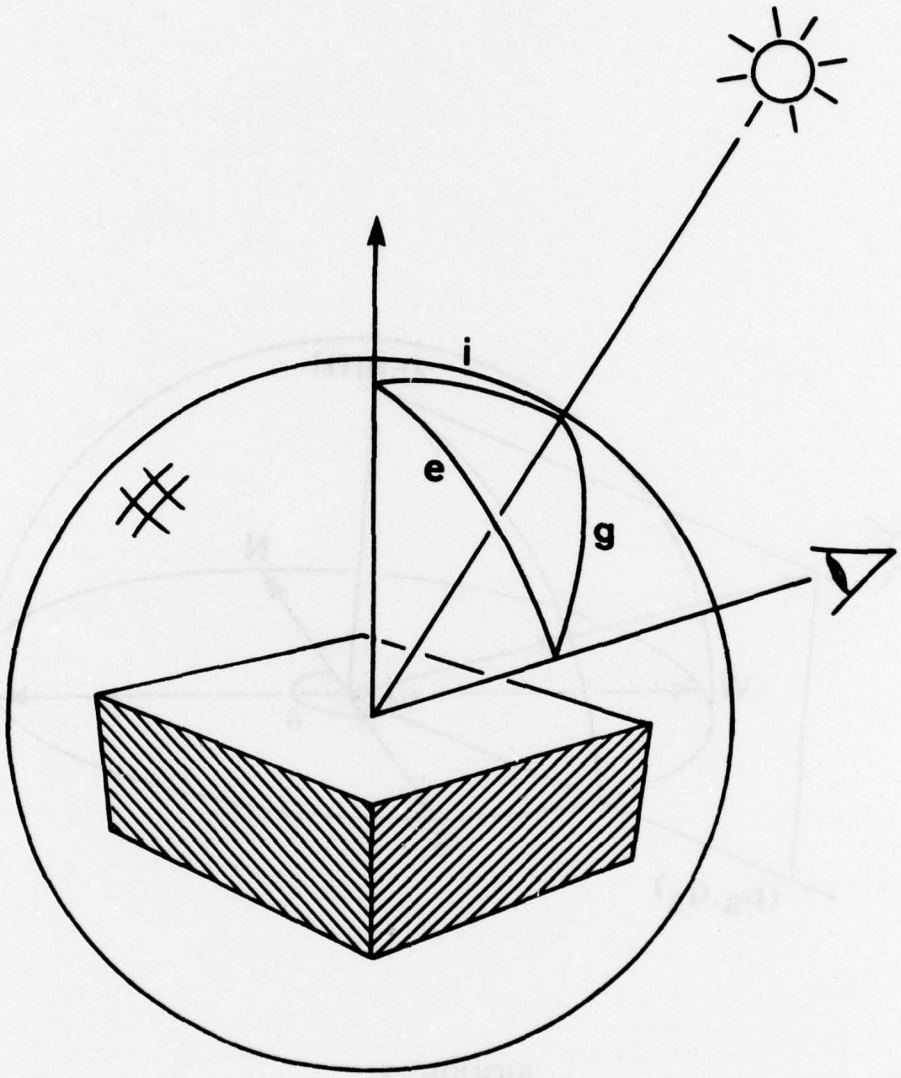


FIGURE 3

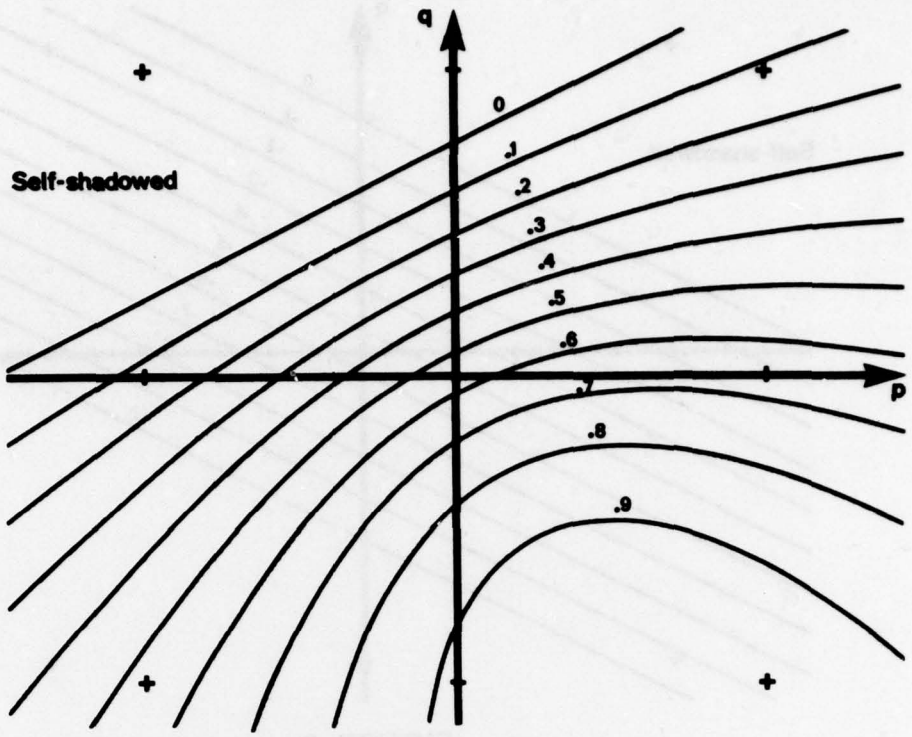
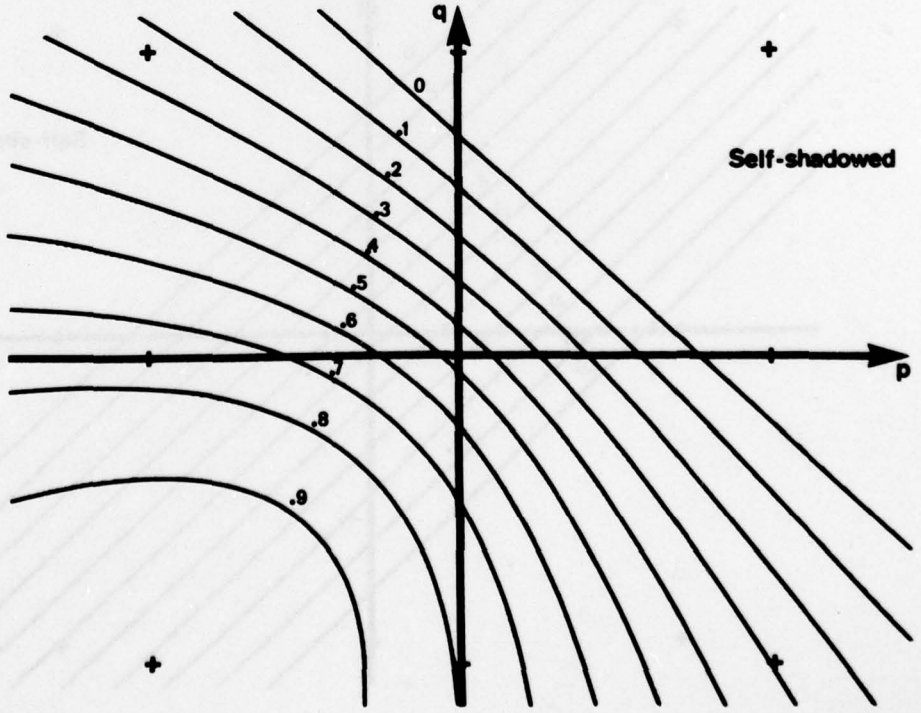


FIGURE 4a

FIGURE 4b



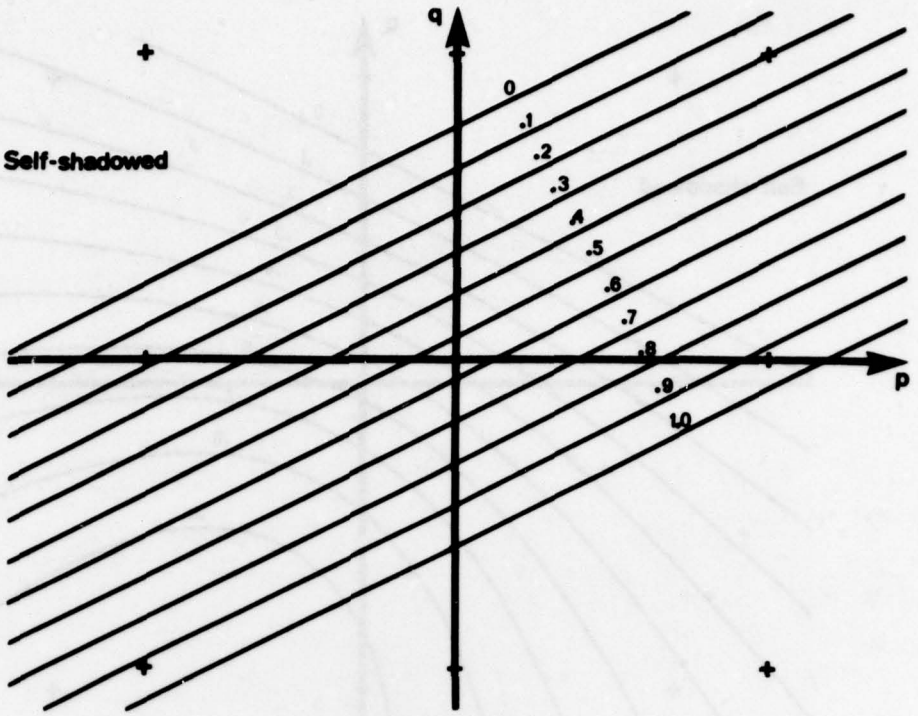
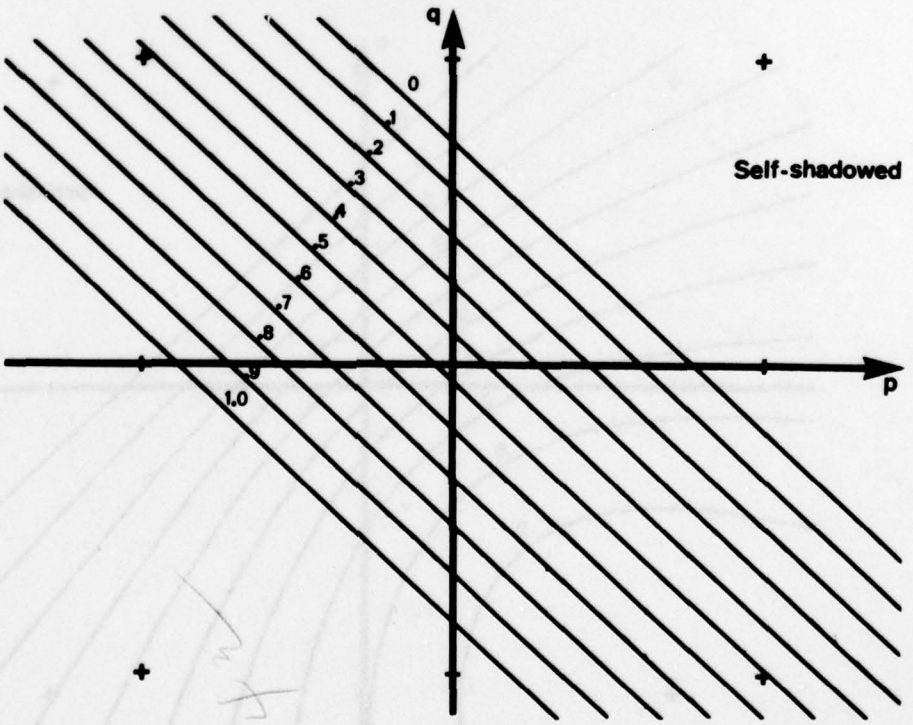


FIGURE 5a

FIGURE 5b



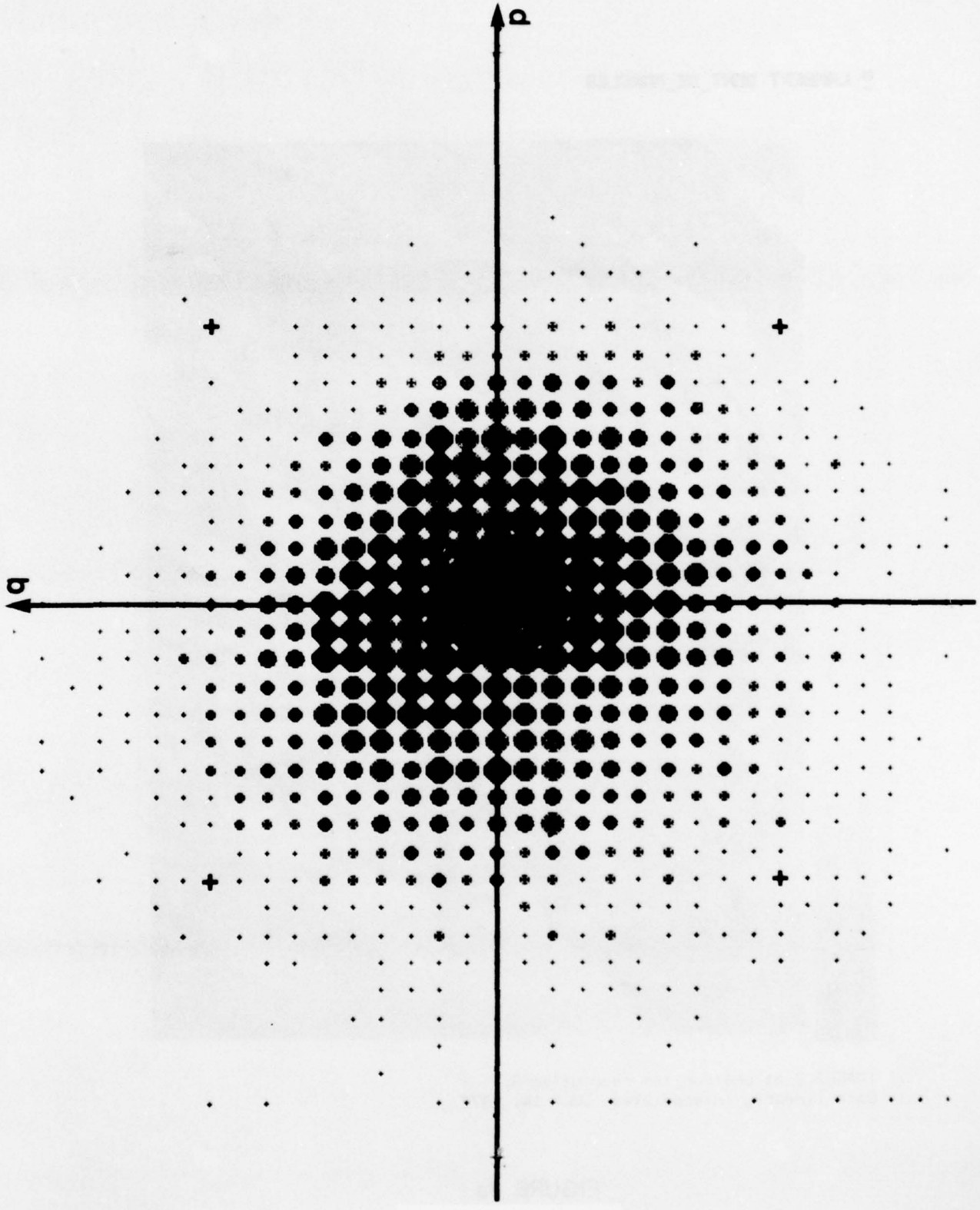


FIGURE 6

LAMBERT DENT_DE_MORCLES

IMAGE2.2 at photowriter resolution 3.
Data linearly interpolated. JULY 14, 1977.

FIGURE 7a

FIGURE 7b

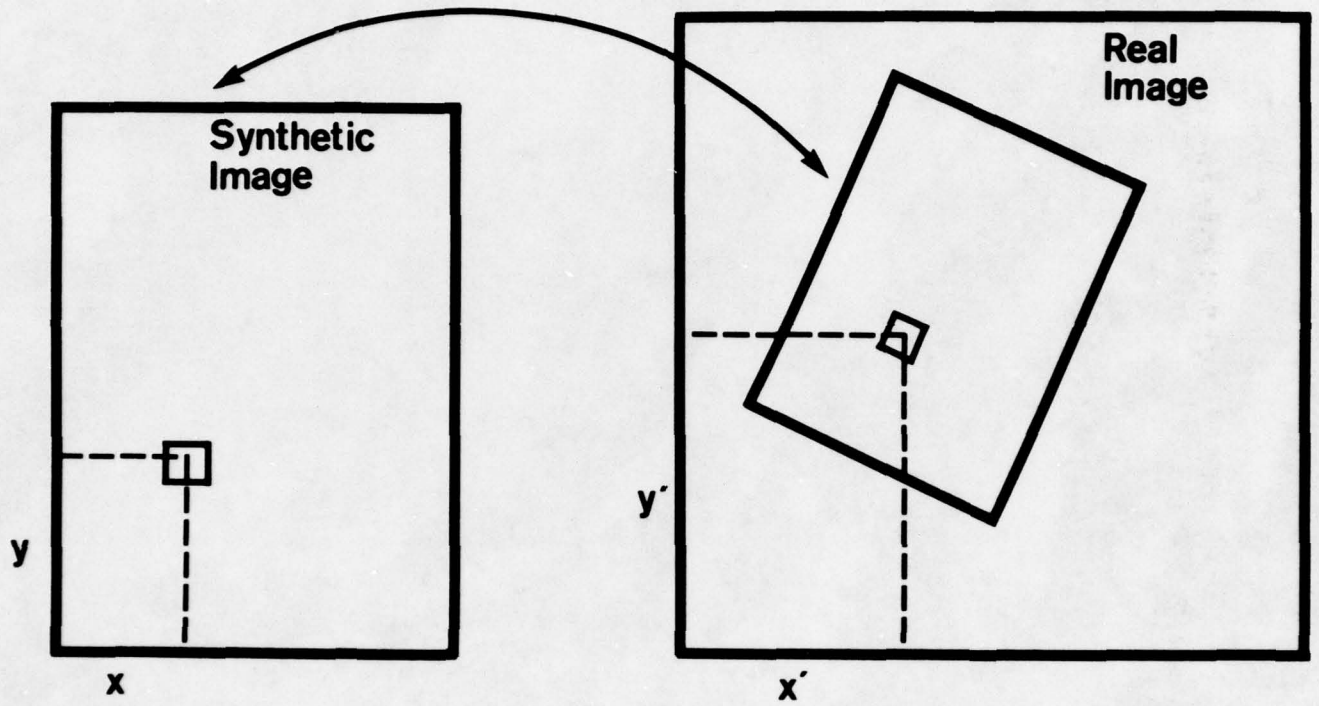


FIGURE 8

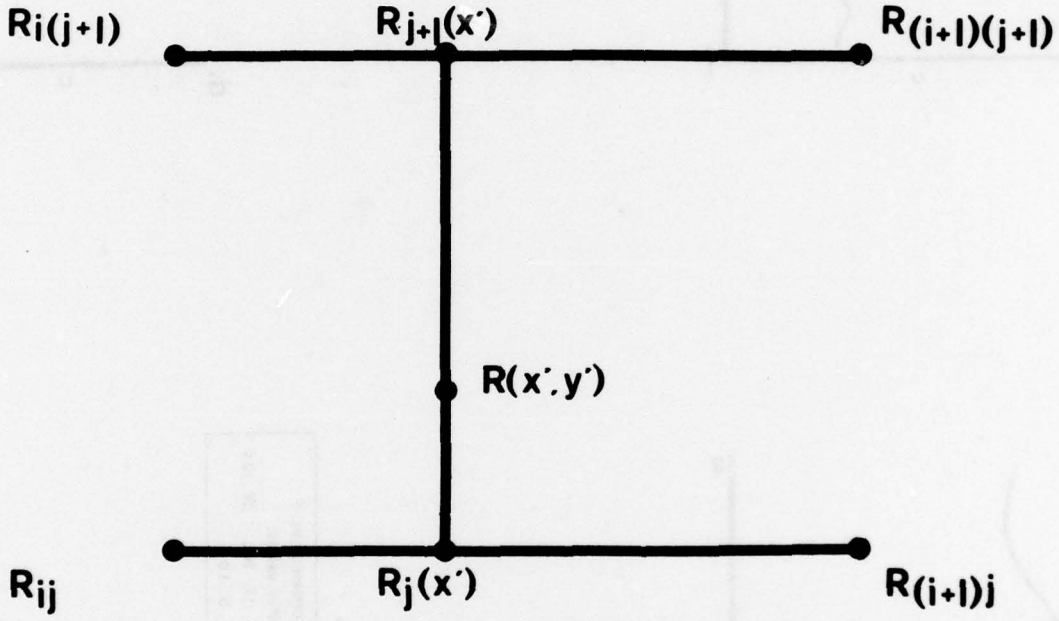


FIGURE 9

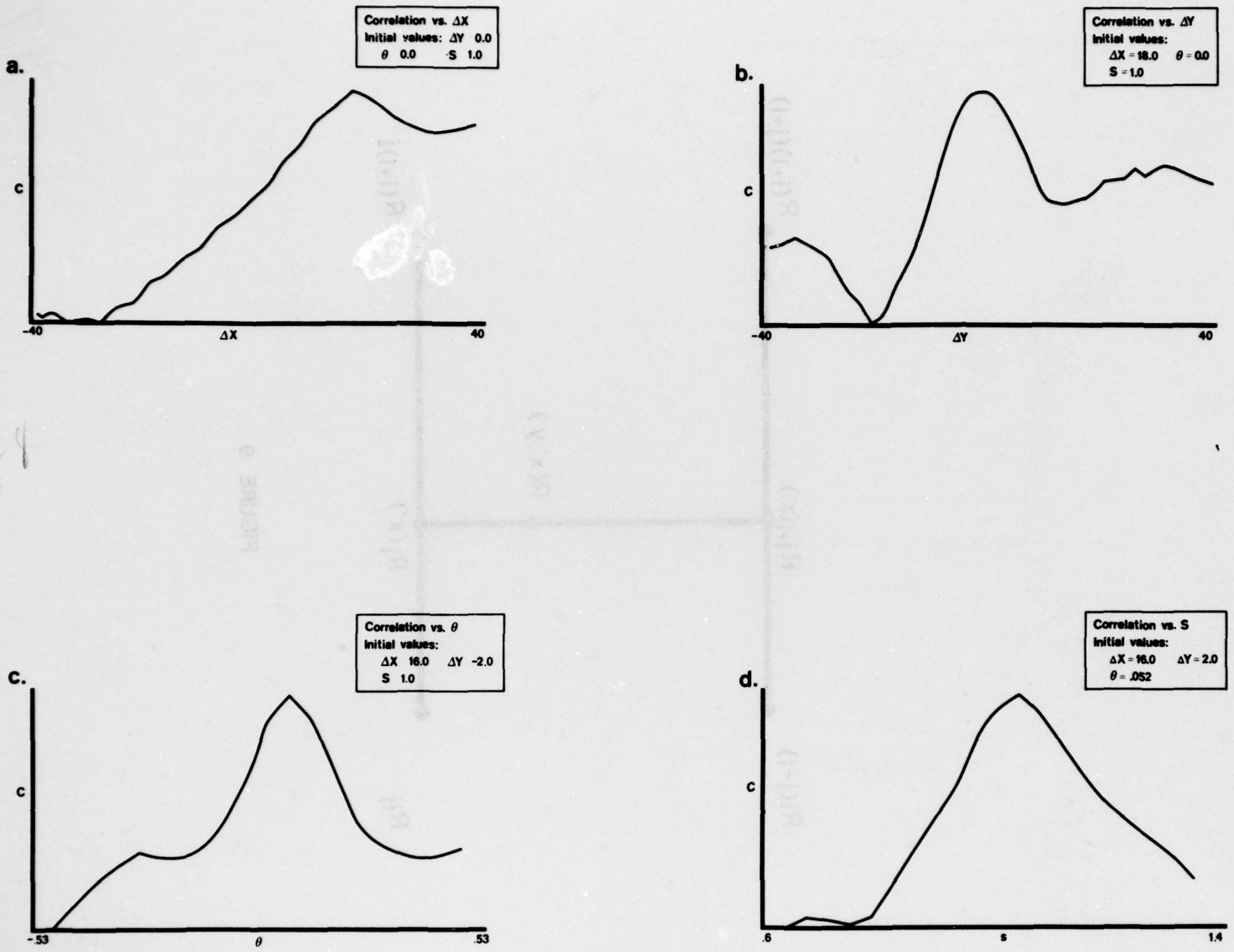

a.
Correlation vs. ΔX
Initial values: ΔY 0.0
θ 0.0 S 1.0
c
-40
ΔX
40
b.
Correlation vs. ΔY
Initial values:
ΔX = 18.0 θ = 0.0
S = 1.0
c
-40
ΔY
40
c.
Correlation vs. θ
Initial values:
ΔX 16.0 ΔY -2.0
S 1.0
c
-.53
θ
.53
d.
Correlation vs. S
Initial values:
ΔX = 16.0 ΔY = 2.0
θ = .052
c
.6
s
1.4

FIGURE 10

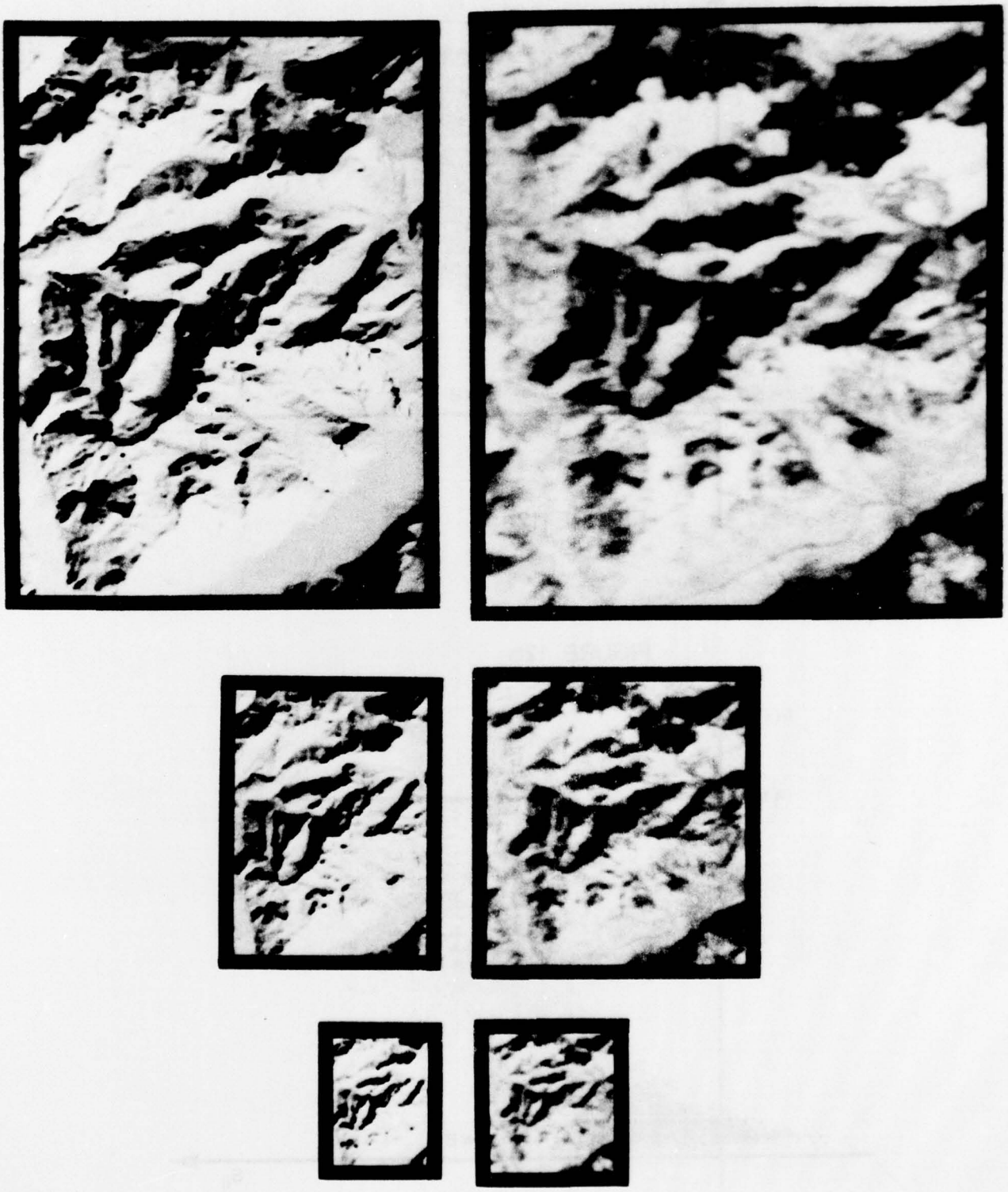


FIGURE 11

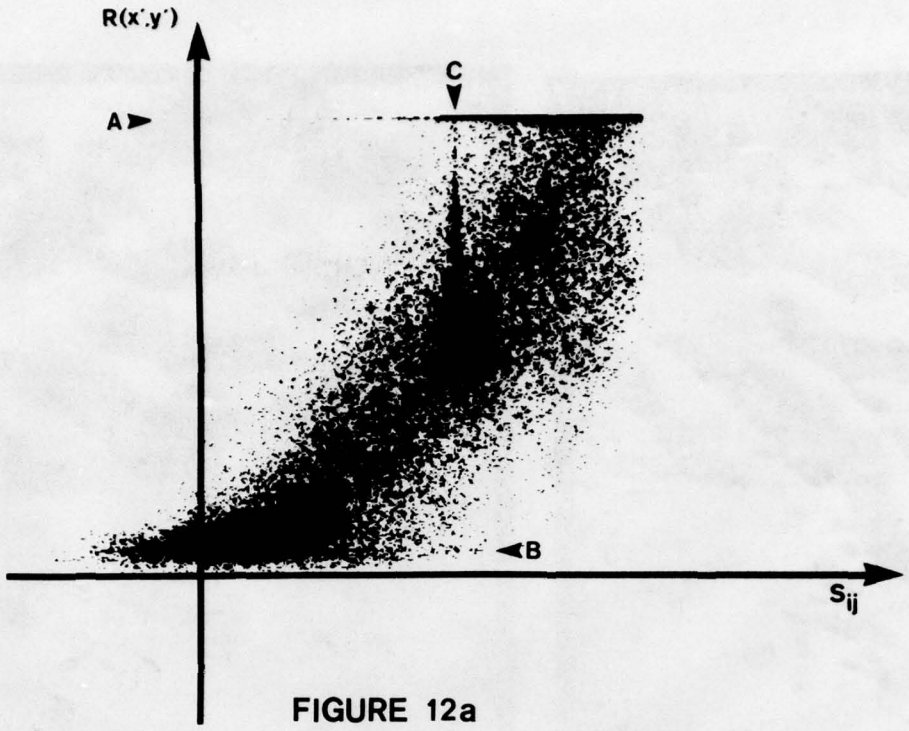
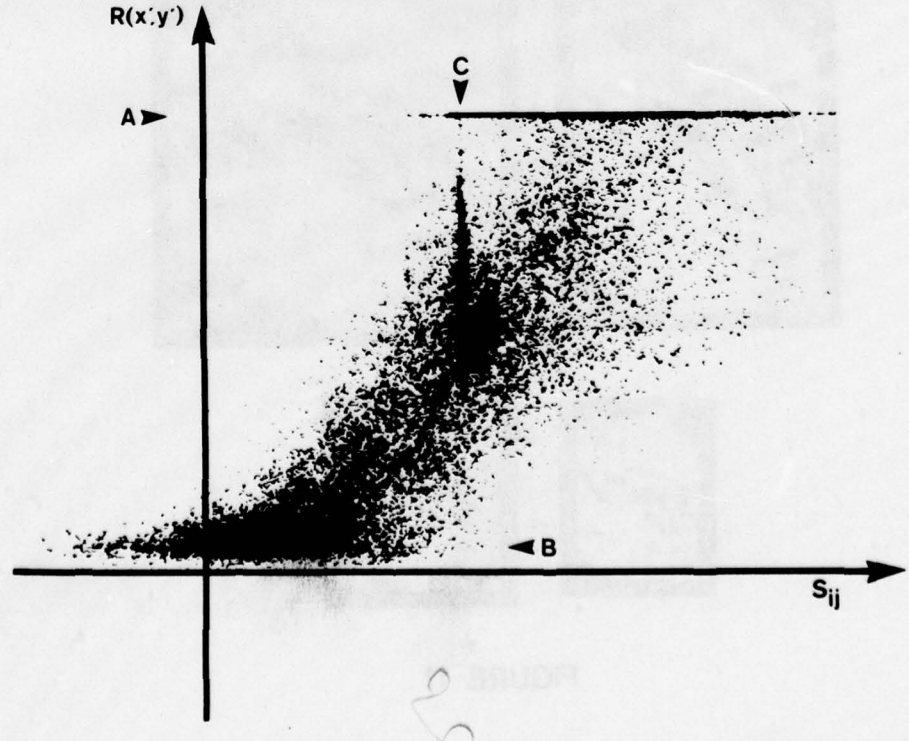
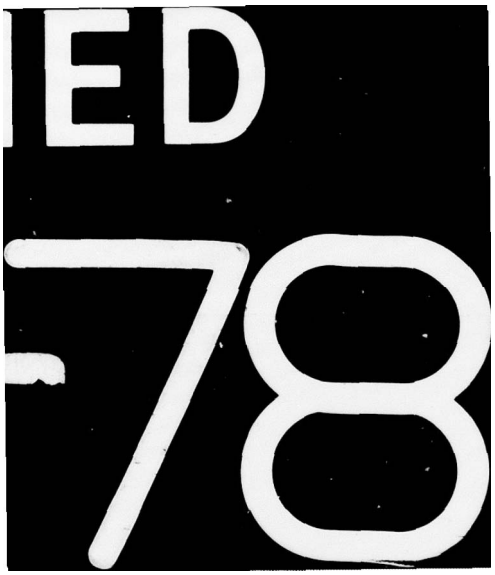


FIGURE 12a

FIGURE 12b




ED
78